Therapeutic Innovation
& Regulatory Science
2014, Vol. 48(6) 749-761

DOI: 10.1177/2168479014527747
tirs.sagepub.com

Implementing Adjusted Imaging Metrics Within Roche With the Metrics Champion Consortium: Experiences and Outcome

Andrew McDonald, BSc[1] and Gudrun Zahlmann, PhD[1]

Abstract

Purpose: To implement adjusted performance imaging metrics on imaging clinical trials of a pharmaceutical company (Roche) in a business relationship with preferred imaging providers and to report on findings and lessons learned. *Methods:* In 2009 the Metrics Champion Consortium provided the first imaging metrics for use in clinical trials as industry consensus. Roche reviewed, adjusted, excluded, and extended these metrics and defined target values per metric in order to implement them in all clinical trials with 7 preferred providers. *Results:* Roche preferred providers were able to report on all 19 metrics (8 unchanged Metrics Champion Consortium, 7 adjusted, and 4 Roche defined). Seventy-three Roche studies over 27 months form the basis for reporting; data are provided as mean and standard deviation per disease area with number of studies and for all studies reported for the specific metric for all providers. Disease areas are oncology, central nervous system, and inflammation. Seventeen metrics have proven to be useful; 2 metrics did not provide sufficient information; and 4 metrics need adjustments of target values. *Limitations:* Imaging trial–related metrics are a new concept, and Roche and providers had to develop the same consistent understanding of content and how to report a specific metric. The 73 studies covered all phases and disease areas, which made it difficult at times to compare results. *Conclusions:* Imaging metrics in clinical trials are a useful tool in improving timeliness and quality of imaging data, enhancing trial processes, and governing sponsor/provider relationships. It increases the transparency in the business relationship and in the different clinical trial–related process steps. The use of metrics highlights critical topics, such as reading and adjudication, and enables parties to take actions to improve performance. Disease area–related reporting needs more data for specific improvements.

Keywords

clinical trials, imaging, performance metrics

Background

Drug development is a high-risk and lengthy process. The pharmaceutical industry tries to reduce the cost of drug development by looking into all process steps regarding potential improvements. Especially for clinical development, the focus is typically on assessing value-critical steps by reviewing business opportunities and potentially changing procedures.[1] One possibility identified is to designate mission-critical process steps as company internal processes and conversely outsource to external partners where they can be more cost-effective while preserving the necessary quality level.

All new drugs must go through clinical testing in order to prove efficacy and patient benefit before they can be marketed. Clinical trials are conducted in a relationship between: clinical sites that manage the patients in the trial, service providers (outsourcing model), and the pharmaceutical company as the trial sponsor with ultimate responsibility. Depending on the phase of development that the clinical trial is supporting, the complexity of the clinical investigation (which questions of clinical development need to be answered), the regulatory relevance, the risk level, and the expected time for results reporting, there are a variety of opinions and models as to what extent to work with service providers. Outsourcing can deal with all tasks mentioned or a selection of one or more.

From a trial sponsor perspective, all outsourcing requires an appropriate business model and business relationship

[1] F. Hoffmann-La Roche Ltd, Basel, Switzerland

Submitted 22-Nov-2013; accepted 18-Feb-2014

Corresponding Author:
Andrew McDonald, F. Hoffmann-La Roche Ltd, GPPS Building 663, U2013, Grenzacherstrasse 124, Basel, 4070 Switzerland.
Email: andrew.mc_donald@roche.com

Table 1. Unchanged MCC metrics with Roche target values.

Metric No.			
Roche	MCC	MCC Metric Description	Roche-Defined Target
11	1	Average percentage of variance in the imaging budget	<20%
10	2	Average no. of calendar days from imaging study award to contract signature	30 d, 100%
1	3	Site start-up: percentage of sites qualified vs actual	80% qualification within 1 mo of site identified being communicated to Imaging Contracts Research Organization (iCRO)
2	5	Image acquisition: average no. of calendar days from image acquisition at clinical site to image receipt at preferred provider	14 d
3	6	Image acquisition: average no. of calendar days from image receipt to initial quality feedback to site	7 d
4	8	Image processing: average no. of calendar days from image receipt to ready for independent review	14 d
17	19	No. of image acquisition technique-related amendments per modality per protocol	$\leq$1 per modality per protocol
15	E1-E4	Reviewer quality: Inter- and intrareader variability	<20% variance from defined target

d, calendar days; MCC, Metrics Champion Consortium.

management, including performance review. To gain the most efficiency, based on experience between service providers and pharmaceutical companies, there is a tendency to form long-term relationships with service providers—so-called preferred provider relationships. Vendors working in a preferred-provider relationship with Roche have agreed to report performance via well-defined metrics. This reporting is on a per-study basis. In cases where more studies are supported by a given preferred provider (henceforth, simply referred to as "providers"), an overview across all studies is requested in addition.

Metrics are not a new concept but are used in many industries. According to a common definition,[2] metrics (or key performance indicators) are "quantifiable measurements, agreed to beforehand, that reflect the critical success factors of an organization." For imaging in clinical trials specifically, there exists the Metrics Champion Consortium (MCC), with its mission "to help sponsor, service provider and clinical site organizations involved in the pharmaceutical, biotechnology and medical device industries improve their overall clinical trial development processes through the utilization of MCC standardized clinical trial performance metrics."[3] MCC worked with various industry stakeholders in 2008 and 2009 on developing and providing those metrics for imaging trials. So far, the measurements to be used are published, but there is no reporting of specific results following their implementation.[4]

At Roche, 90% of all oncology clinical trials have an imaging component. Central nervous system (CNS) and inflammatory disease trials also use imaging but to a lesser extent. Therefore, imaging-related trial processes are an attractive target for improving overall clinical trial processes. Roche selected 7 preferred imaging providers in 2009. As part of the governance process, provider relationship management metrics for imaging clinical trials were defined, agreed with the providers, and implemented.

This article outlines the Roche imaging-specific metrics. It then offers the results achieved by applying those metrics to 73 clinical trials over 9 quarters, from fourth quarter 2010 to fourth quarter 2012, in working with Roche's 7 imaging providers. Finally, it discusses the usefulness of the metric concept and individual metrics in application to study conduct.

Materials and Methods

The aim was to design and implement a process and a set of parameters to measure performance for governance and study progress. In addition, the imposition of a metric structure was intended to build transparency into the process and to enable Roche to spot trends earlier than what was previously possible.

The MCC imaging metrics v. 1.0—19 defined metrics and 1 exploratory metric on reader quality (inter- and intrareader variability)—were published in 2009 and form the basis of the metrics collected by Roche and used in the governance reports.[3] The metrics selected by Roche for imaging clinical trials are the following: metrics for assessing study progress, metrics for assessing provider performance for governance, and metrics for collecting data for benchmarking purposes. In addition to the MCC's metrics definition, Roche defined quantifiable targets for each metric based on feedback from all providers and internal benchmarks. Roche selected 14 of the 20 MCC imaging metrics for the governance report, as these were

Table 2. Adjusted MCC Metrics With Roche Target Values.

Metric No.			
Roche	MCC	MCC Metric Description	Roche Target: Interpretation and Adjustments[a]
5	9	Image processing: average no. of calendar days from when the image is designated for review to completion of the review	*14 d:* No. of days, "Actual date ready for reading" vs "Actual date radiology read complete" *21 d:* No. of days, "Actual date ready for reading" vs "Date adjudication complete" *28 d:* No. of days, "Actual date ready for reading" vs "Actual date oncology / cardiology review complete"
6	13	Percentages of missing imaging visits	*<5%:* Percentage of missing scans, "Expected image date" vs "Actual image date" *14 d:* No. of days when missing scans are identified, from "Date query sent" to "Date returned"
8	15	Image queries: Average no. of calendar days an imaging query is outstanding	*7 d:* No. of days, "Date query raised" to "Date query sent" *14 d:* No. of days, "Date query sent" to "Date returned"
7	16 + 17	Export submission: Average no. of calendar days from last patient reviewed to delivery of results dataset to Roche	*<30 d:* No. of days, "Actual date radiology read complete" or "Date adjudication complete" or "Actual date oncology / cardiology review complete" (dependent on study, whichever is last date) vs "Date scan sent to Roche (or archive)" *<5 d:* No. of days, "Expected date scan clean and ready for archiving" vs "Actual date scan clean and ready for archiving"
16	18	Independent review charter lifecycle	*42 d:* No. of calendar days to produce first draft independent review charter, from "Date contract signed" to "Date first draft charter" *60 d:* No. of calendar days to finalize independent review charter, from "Date first draft charter" to "Date charter final" *4 and 0:* No. of versions, "No. of drafts prefinalzation" and "No. of changes postfinalization" *60 d:* No. of calendar days to finalize reading database, from "Date charter final" to "Reading database final"

d, calendar days; MCC, Metrics Champion Consortium.
[a]Adding granularity by adding submetrics.

deemed most appropriate to measure provider performance; however, some metrics could also be used to track site or Roche performance. Eight of the metrics used are from the MCC with no changes (Table 1). Certain MCC metrics were adjusted by Roche. as shown in Table 2.

The Roche adjustments above were mainly refinements to MCC metrics and definitions to provide better clarity. For example, MCC metric 18 is ''Number of weeks to develop and write independent review charter.'' Roche has clarified this further as follows: number of days to produce first draft independent review charter, number of days to finalize independent review charter, number of drafts prefinalization, number of changes postfinalization, and number of days to finalize reading database after charter finalization. Roche is using these results to actively contribute to discussions on updating the MCC imaging metrics in a second version.

Six additional metrics are Roche defined and described in Table 3. The first Roche-defined metric is a quality assessment for Roche purposes; the second describes stability of project staffing. Since reader performance and adjudication quality are of specific importance, those metrics were added too.

The following MCC metrics were not used:

MCC metric 4: site start-up (days from site designated ready to first date of image receipt). Roche had set the target for the other site start-up metric (MCC metric 3, percentage of sites qualified vs actual) as a percentage within a designated period.

MCC metric 7: days from image quality control complete to reporting of eligibility results. The measure is study specific and tracked by Roche in study-specific reports.

MCC metrics 10-12: image quality. Metrics are intended to measure site performance rather than provider performance.

MCC metric 14: percentage of site queries. Metric measures site performance rather than provider performance.

MCC metric 17: days from original estimate to actual for export submission. These data are tracked in study-specific reports, and Roche also thought that MCC

Table 3. Roche-defined metrics.

Metric			
No.	Description	Roche Target	Clarification
9	Audit findings, no. of critical findings	0 critical	If an audit has been completed in a specific quarter to be reported in the overall average metric tab, otherwise to be left blank.
12	Turnover of key project staff (project manager, scientific team lead, data team lead, and readers)	1 per study per year	Used as a risk indicator to check change management (succession planning) is conducted adequately enough such that staff turnover will have minimal impact on a study.
13	Reviewer status: percentage variance adjudication, expected vs actual	<5% increase	The adjudication rate in charter or contract vs actual. It is expected that a projected adjudication rate is to be entered into each charter; this metric will then track the variance in expected vs actual in each quarter for each study when reading starts. This enables Roche to be specific regarding expectations on reviewer agreement for a reading task and reflects complexity of read instead of using 1 target number for all reads.
14	Reviewer status: percentage difference in expected vs actual reads	<10% for all	Percentage: (1) "Patients ready for review, expected" vs "Patients ready for review, actual"; (2) "Patients read, expected" vs "Patients read, actual"; (3) "Patients queued for adjudication, expected" vs "Patients adjudicated, actual"; (4) "Patients adjudicated, expected" vs "Patients adjudicated, actual"; (5) same as 1 and 2 but for "Cumulative oncology/cardiology ... review, expected vs actual." This is to be reported for all data in a quarter for each study and is designed to show the provider's forecasting vs actual capabilities.
18	Financial health of provider	<0.5%	Bankruptcy prediction score, percentage; 2× per year
19	Accurate quarterly forecasting	<3%	Percentage difference between actual spend and forecast spend.

metric 16 (days from last patient reviewed to delivery of data set) was more appropriate.

A formal governance process enables teams to deliver the contracted work (1) by leveraging resources, infrastructure, process improvements, and any learning across programs and functions and (2) by providing oversight and management of operational delivery through metric monitoring, pipeline and capacity planning, performance management, process/standards harmonization, continuous improvement, issue escalation, and planning of future collaborations. It also maintains a focus on the major objectives and opportunities of the relationship and helps strategic planning.

The governance committee in Roche for imaging providers uses the above metrics to ensure that both parties are informed of the overall status of each project and any issues. Roche uses a traffic-light system based on the method shown in Table 4.

The metrics are reported quarterly by the providers to Roche in 2 forms. The first is a summary overview on all studies, and the second contains each study attached as a separate table so that governance team members will be able to investigate any studies where a provider is not reaching a specific target. Each study metric sheet will also be sent by the provider's project manager to the respective Roche study team.

Table 4. Traffic-light system.[a]

Green	Meets target
Orange	Does not meet target ≤10%
Red	Does not meet target >10% (action plan to be attached)
Blue	Does not meet target but, due to specific requirements of a study, is not an issue (reason[s] to be included)

[a]Color available online only.

Roche organizes quarterly governance meetings with all providers. The reported metrics form the basis for status review and discussion on improvements. Each quarter, the metrics for each provider are combined with other providers for comparison. The metrics are also collected and analyzed for each quarter sequentially to see how a specific metric changes over time for a specific provider (due to new studies or to process improvements).

In the first four quarters reports created under this scheme, the metrics were captured and reevaluated for all parties to get acquainted with the reporting process and to ensure consistent reporting from all providers. In the subsequent quarters, apparent trends are being considered if the data have sufficient meaning and/or statistical power (eg, number of data points, number of reads). For example, reporting noncompliance on the adjudication rate metric (metric 15) in a single study occurrence will

Table 5. Roche metric 1: percentage of sites qualified vs total number of sites.

Metric #	Metric Description	Target	ALL					CNS			ONCOLOGY			INFLAMMATION		
			Mean	SD	min	max	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
1	Site start-up: Percentage of sites qualified vs total number of sites	80% qualification within 1 month of site identified being communicated to iCRO	77.0%	30.6%	0.0%	100.0%	49	68.1%	28.7%	7	78.8%	31.9%	38	75.7%	23.9%	4

Table 6. Roche metric 2: time from image acquisition at site to receipt at preferred provider.

Metric #	Metric Description	Target	ALL					CNS			ONCOLOGY			INFLAMMATION		
			Mean	SD	min	max	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
2	Image Acquisition: Average number of calendar days from image acquisition to image receipt (at CRO)	14 d	22.8	17.4	2	86	64	3.6	1.3	7	27.3	17.3	49	12.4	7.8	8
		electronic	17.1	9.6	7	27	5	—	—	—	19.4	9.3	4	7.9	—	1
		courier	26.7	13.5	8	41	5	—	—	—	29.1	14.3	4	16.9	—	1

result in an immediate discussion on this specific topic, will be followed up in the next governance meeting, and, if necessary, will be included in topics of concern for future audits. Other cases, such as the second time that an orange reading is reported for data transfer from clinical site to provider (metric 2), will be followed in governance discussions, but no immediate action will be taken if the study goal is not affected. Thereby, with such reporting and review, there exists an escalation process in the event of nonperformance, enabling transparent decision processes for the overall relationship between Roche and provider.

Results

For 73 clinical trials (58 in oncology, 7 in CNS diseases, and 8 in inflammation), data from 9 quarters (27 months) are analyzed herein. Means and standard deviations of all parameters and metrics from all quarters, with data per metric for each study, were fed into overall and disease area–stratified results. These measures were used to get an overall impression on metric mean and variability and to support an assessment of usefulness and validity of introduced target values. The number of studies included in each metric analysis is outlined here per metric and disease area. Overall, 73 studies are reported; however, the numbers of studies differ per specific metric analysis, depending on the available data. Since the sample sizes for studies per metric and disease area are very small, further statistical analysis at higher detail has not been performed. Results are reported in numerical order on the basis of Roche metrics numbering.

Metric 1: Percentage of Sites Qualified vs Total Number of Sites (Table 5)

Metric 1 shows that all 49 reported studies are just below target and that there is a marked difference in the results dependent on the disease area, with most studies being reported from oncology (n = 38). The standard deviation shows limited variability in the data. CNS studies show slower site qualification than do oncology and inflammation studies.

Metric 2: Time From Image Acquisition at Site to Receipt at Provider (Table 6)

Metric 2 shows clear disease-area dependence. Oncology trials are not meeting the target, whereas inflammation and CNS studies do. This metric is highly dependent on site responsiveness. Data for electronic versus courier image delivery confirm that electronic delivery is faster and can contribute to meeting the target. However, the number of studies reported in this metric is very small.

Metric 3: Time From Image Receipt at Provider to First Feedback to Site (Table 7)

Metric 3 is important in that it provides feedback to the clinical site on the quality of its images in good time to ensure that issues are put right prior to future scans. There is a clear difference among disease areas, with oncology being most off-target. The 2 CNS providers clearly organize immediate feedback to sites, whereas in oncology, with 6 providers, there is room for improvement. High standard deviations for oncology and inflammation illustrate a high level of compliance variability among studies. The model of reading employed for a specific study might also influence this metric.

Metric 4: Time From Image Receipt to Being Ready to Read (Table 8)

Metric 4 includes all activities needed by the provider in working with sites to get good quality images via queries for the

Table 7. Roche metric 3: time between image receipt at preferred provider to first feedback to site.

Metric #	Metric Description	Target	ALL					CNS			ONCOLOGY			INFLAMMATION		
			Mean	SD	min	max	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
3	Image Acquisition: Average number of calendar days from image receipt to initial quality feedback to site	7 d	12.4	11.9	0	60	66	0.8	0.7	7	14.8	12.4	51	7.4	4.2	8

Table 8. Roche metric 4: time from image receipt to being ready to read.

Metric #	Metric Description	Target	ALL					CNS			ONCOLOGY			INFLAMMATION		
			Mean	SD	min	max	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
4	Image Processing: Average number of calendar days from image receipt to ready for independent review	14 d	28.5	31.7	0	172	55	1.0	1.1	7	33.7	33.5	42	23.7	18.3	6

provider to be ready for reading. For all studies, this metric is out of target. Again, there is a clear disease-area difference. The 1-day mean for CNS is provider specific: whereas the CNS vendor has processes and systems in place to enable read readiness within a day, the other provider for 1 study takes a mean of 3 days to accomplish this. However, oncology studies need more than double the target time and skew other disease-area results. This reflects, in part, batch-reading procedures for oncology studies.

Metric 5: Time From Images Being Ready to Read to Read Being Completed (Table 9)

Metric 5 describes the quality of the provider in managing the central reading process. Provider processes are apparently ensuring adequate target delivery here.

Metric 6: Percentages of Missing Imaging Visits (Table 10)

Under metric 6, submetric 6.1 is the amount of missing images at a specific time point, whereas submetric 6.2 is the amount of time taken to resolve the query once the clinical site is informed. By including both submetrics, we try to understand what the reason for any missing scans might be. Our results show that, on average, 10% of images are missing at each quarter, with most coming from PET-CT (positron emission tomography–computed tomography), RECIST, and Cheson studies. As expected, one reason for missing data is communication with clinical sites, which needs to be improved, as results for submetric 2 show.

Metric 7: Export Submission—Days From Last Patient Reviewed to Delivery of Data Set (Table 11)

In metric 7, the mean value shows all targets are being met. Submetric 7.1 reporting is far below the target value; hence, the target might need adjustment.

Metric 8: Image Query Turnaround Times (Table 12)

Submetric 8.1 reflects internal provider processes, whereas submetric 8.2 reflects communication of providers with clinical sites. Again, there is a clear difference between CNS and other disease areas in site communication.

Metric 9: Audit Findings (Table 13)

Audits are conducted on specific studies or as a general capability; the aim is to have no critical (most serious) findings. In the 10 audits conducted in the last 9 quarters, only 1 resulted in 1 critical finding.

Metric 10: Time From Study Award to Contract Signature (Table 14)

Metric 10 shows a mean of 63 days to put a contract in place from formal award of the study to contract signature and is influenced by both Roche and provider processes. However, this metric is being reduced with the implementation of set contractual templates and processes; the last 12 months (first quarter to fourth quarter 2012) shows a mean of 24 days to put a contract in place.

Metric 11: Variance in Imaging Budget (Table 15)

There is a standard, unitized cost structure in place with all Roche providers. Metric 11 measures the variance from the original budget to any new budget documented in additional work orders. Our results show that providers largely adhere to this unitized cost structure, whereas the variance is due to changes in study assumptions only (timelines, number patients, etc). Due to the reduction in scope on several studies, oncology shows a negative mean, whereas CNS shows a result above target, due to increases in assumptions (doubling of patients and timelines).

Table 9. Roche metric 5: time from images being ready to read to read being completed.

Metric #	Metric Description	Target	Clarification	ALL					CNS			ONCOLOGY			INFLAMMATION		
				Mean	SD	min	max	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
5	Image Processing: Average number of calendar days from when the image is designated for review to completion of the review	14 d	1) # days "Actual Pending Final Review Date" vs "Actual Read Completion Date"	12.2	11.4	0	46	41	4.0	—	1	12.0	11.8	35	15.7	8.9	5
		21 d	*2) # days "Actual Pending Final Review Date" vs "Adjudication Competion Date"*	13.6	15.5	0	43	11	0.4	—	1	11.7	12.9	9	43.2	—	1
		28 d	3) # days in "Actual Pending Final Review Date" vs "Actual Date Oncology/ Cardiology... review complete"	11.8	10.2	1	32	15	—	—	—	11.9	10.6	14	10.6	—	1

Table 10. Roche metric 6: percentages of missing imaging visits.

Metric #	Metric Description	Target	Clarification	ALL					CNS			ONCOLOGY			INFLAMMATION		
				Mean	SD	min	max	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
6	Percentages of missing imaging visits	<5%	1) % of missing scans identified "Anticipated Exam Date" vs "Actual Exam Date"	10.0%	18.3%	0.0%	100.0%	67	0.3%	0.4%	7	11.9%	20.2%	52	6.2%	6.1%	8
		14 d	2) # days from when missing scans are identified from "Date Query Issued" to "Date Query Resolved"	23.1	26.9	4	152	44	11.0	5.0	4	21.4	25.4	35	43.9	39.6	5

Table 11. Roche metric 7: export submission—average number of calendar days from last patient reviewed to delivery of data set.

Metric #	Metric Description	Target	Clarification	ALL					CNS			ONCOLOGY			INFLAMMATION		
				Mean	SD	min	max	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
7	Export Submission: Average number of calendar days from last patient reviewed to delivery of dataset	<30 d	1) "Actual Date radiology read complete" vs "Date scan sent to Roche (or archive)"	9.3	8.6	0	30	13	—	—	—	6.5	6.3	9	15.8	10.4	4
		<5 d	2) "Expected Date scan clean and ready for archiving" vs "Actual Date scan clean and ready for archiving"	4.9	8.4	0	18	4	—	—	—	6.2	9.8	3	1.3	—	1

Metric 12: Staff Turnover (Table 16)

Metric 12 was introduced based on previous experience, as it can identify specific providers where high turnover affects performance and study team satisfaction. All studies met target value.

Metric 13: Adjudication Rate (Table 17)

Metric 13 shows that CNS and inflammation studies are tracking to target, with oncology over target, with 11.1% mean variance and high standard deviation.

Table 12. Roche metric 8: image query turnaround times.

Metric #	Metric Description	Target	Clarification	ALL					CNS			ONCOLOGY			INFLAMMATION		
				Mean	SD	min	max	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
8	Image queries: Average number of calendar days an imaging query is outstanding	7 d	1) Date Query Raised to "Date Query Sent"	2.9	4.6	0	23	56	0.0	0.1	7	3.3	4.9	41	3.6	3.9	8
		14 d	2) Date Query Sent to "Date Returned"	19.2	14.4	4	85	53	6.8	2.2	7	20.6	14.8	39	23.6	14.2	7

Table 13. Roche metric 9: audit findings.

Metric #	Metric Description	Target	ALL					CNS			ONCOLOGY			INFLAMMATION		
			Mean	SD	min	max	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
9	Audit Findings, number of critical findings	0 critical	0.1	0.1	0	1	10	0.0	—	1	1.0	1.0	8	0.0	0.0	1

Table 14. Roche metric 10: time from study award to contract signature.

Metric #	Metric Description	Target	ALL					CNS			ONCOLOGY			INFLAMMATION		
			Mean	SD	min	max	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
10	Average number of calendar days from imaging study award to contract signature	30 d	63.0	42.1	12	180	34	90.0	—	1	56.5	37.2	30	118.7	60.0	3

Table 15. Roche metric 11: variance in imaging budget.

Metric #	Metric Description	Target	ALL					CNS			ONCOLOGY			INFLAMMATION		
			Mean	SD	min	max	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
11	Average percentage of variance in the imaging budget	<20%	4.2%	26.4%	−62.3%	86.0%	18	43.0%	60.8%	2	−2.3%	20.2%	12	4.4%	5.7%	4

Metric 14: Reviewer Status Expected vs Actual Reads (Table 18)

Metric 14 describes the planning process for reading and adjudication versus actual data. Oncology studies exceed 3 times the target value for read planning (submetric 14.1). All other disease areas are on target.

Metric 15: Inter- and Intrareads Variability (Table 19)

Metric 15 has been reported as a sum of all reader variability to date. It shows clear differences according to disease area, with inflammation exceeding the target by more than double the target value.

Metric 16: Imaging Charter Development Lifecycle (Table 20)

Metric 16 describes the important charter-writing process, and it has three parts. The first (submetric 16.1) describes the providers' charter-drafting process and shows that, other than studies in inflammation, all other charter writing takes longer than the target number of days. In CNS, the charter-drafting process has to be changed and actively monitored. Part 2 (submetrics 16.2-16.4) describes the process of charter review by Roche study teams and communication with providers. Excluding a single CNS study, the number of review cycles or charter cycles meets the target value, but overall, the review process takes too long. Postfinalization changes are reported in only 12 studies in oncology and 1 study in inflammation. Part 3 reflects the charter translation into an actionable reading database and system by providers. On average, this process takes too long in oncology.

Metric 17: Changes to Imaging Site Manual (Table 21)

Metric 17 is an MCC metric designed to understand robustness of the site imaging acquisition protocol; most Roche studies are tracking to target. Only inflammation studies are slightly over target.

Table 16. Roche metric 12: staff turnover.

Metric #	Metric Description	Target	ALL					CNS			ONCOLOGY			INFLAMMATION		
			Mean	SD	min	max	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
12	Turnover of Key Project Staff (project manager, scientific team lead, data team lead and readers)	1 per study/year	0.7	0.9	0	4	60	0.4	0.8	7	0.8	0.9	45	0.4	0.5	8

Table 17. Roche metric 13: adjudication rate.

Metric #	Metric Description	Target	Clarification	ALL					CNS			ONCOLOGY			INFLAMMATION		
				Mean	SD	min	max	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
13	Reviewer Status: % variance adjudication expected vs actual	<5% increase	The adjudication rate in charter or contract vs actual	11.1%	24.6%	–54.6%	52.9%	21	0.0%	0.0%	6	18.9%	30.6%	12	2.2%	8.0%	3

Metric 18: Financial Health of Providers (Table 22)

Metric 18 was a Roche metric designed to show the financial health of each provider. As a robust financial health risk mitigation plan is now in place for each provider, this metric was removed in 2011.

Metric 19: Provider Budget Forecasting (Table 23)

Metric 19 was a Roche metric designed to show the financial forecasting ability on each study. As most Roche teams now conduct a straight-line forecasting and accrual process, this metric was removed in 2011.

Discussion and Limitations

After 27 months of metrics use to assess imaging processes in clinical trials in collaboration among providers, clinical sites, and Roche, the overall experience has been positive. For the assessment of provider performance as used for governance, the traffic-light system works well, as it clearly shows at a quick glance any areas that require further investigation. However, study-specific metrics should always be used to understand why a metric is out of range. As demonstrated, all metrics defined herein can be reported by providers. In the reporting time frame, providers have established the appropriate information technology solutions for near-automatic reporting.

As results of this performance review, metrics can be categorized into 3 groups:

Group 1: Metrics that have proven to be useful for measuring business behavior by respective providers and will remain in use unchanged (metrics 1, 3, 6-9, 11, 12, 14, 16).

Group 2: metrics that have proven to be useful but need refinements (metrics 2, 4, 5, 10, 13-15).

Group 3: Metrics that have shown no value for Roche's business relationships and will be removed from the list (metrics 17-19).

Group 1: Metrics That Have Proven Value and Remain Unchanged

To have clinical sites qualified before the first patient scan is performed is a very important achievement (metric 1). However, this target has not been fully met in the studies under investigation. In part, this reflects the need for process improvements of the providers working with clinical sites. This reflects the ongoing change, especially for early-phase studies, toward a more rigid Roche requirement to ensure high quality of image data for accurate, reliable, and reproducible quantitative imaging results. More and more phantom measurements are requested, and this is clearly a novel challenge for some providers, as well as clinical sites, to be completed within the target time. This will be worked on in future. Initiatives such as the Quantitative Imaging Biomarker Alliance[5] might help to achieve this faster on an international scale.

Metric 3 (days from image receipt to initial quality feedback to site) shows that providers do understand the need to immediately react to any query regarding provided image data sets. The target of 7 days reflects the importance of this metric, but on average, providers' internal quality control processes take too long in our opinion. There is a difference seen between prospective studies (mean of 12 days) and those that collect and hold for later analysis (mean of 19 days), which demonstrates that collect-and-hold studies provide no incentive to perform

Table 18. Roche metric 14: reviewer status expected vs actual reads.

Metric #	Metric Description	Target	Clarification	ALL					CNS			ONCOLOGY			INFLAMMATION		
				Mean	SD	min	max	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
14	Reviewer Status: % difference in expected vs actual reads	<10%	1) "Patients Ready for Review - Expected" vs "Patients Ready for Review - Actual"	20.6 %	75.8 %	75.0%	465.0 %	40	3.3%	8.6 %	7	26.8 %	87.0 %	30	0.0%	0.1%	3
		<10%	2) "Patients Read - Expected" vs "Patients Read - Actual"	1.6%	6.1%	-3.9%	27.4 %	23	—	—	—	1.7%	6.5%	20	0.8%	1.4%	3
		<10%	3)"Patients Queued for Adjudication - Expected" vs "Patients Adjudicated - Actual"	0.9%	3.6%	0.0%	15.0 %	17	—	—	—	0.0%	0.1%	15	7.5%	10.6 %	2
		<10%	4) "Patients Adjudicated - Expected" vs "Patients Adjudicated - Actual"	4.6%	7.4%	0.0%	20.0 %	15	—	—	—	5.0%	7.6%	14	0.0%	—	1
		<10%	5) Same as 14a and 14b but for "Cummulative Oncology/Cardiology... Review Expected vs Actual"	0.0%	0.0%	0.0%	0.0%	2	—	—	—	0.0%	0.0%	2	—	—	—

Table 19. Roche metric 15: inter- and intrareads variability.

Metric #	Metric Description	Target	ALL					CNS			ONCOLOGY			INFLAMMATION		
			Mean	SD	min	max	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
15	Reviewer Quality – Inter-reader and Intra-reader variability	<20%	22.7%	29.9%	0.0%	75.0%	11	5.1%	12.4%	6	24.6%	24.8%	3	73.0%	2.8%	2

a fast query resolution. For studies with a busy imaging schedule or those requiring just-in-time imaging reporting of results (eg, for eligibility), immediate reaction is necessary and achievable. For metric 6 (percentage of missing images), oncology and inflammation studies show weaknesses regarding image completeness, whereas CNS studies largely met the Roche-defined target. As imaging contributes often to primary and secondary study endpoints, data completeness is essential.

The response time by clinical sites to queries regarding missing data sets is rather protracted. It is expected that providers improve their communication with clinical sites to meet the agreed target of 14 days for this metric. This is also supported by metric 8.2. (number of days that a query is outstanding). Yet, these results—when compared to metric 7 (days from last patient reviewed to delivery of results data set to Roche), metric 8.1 (days that an imaging query is raised at provider prior to sending to site), and metric 9 (critical audit findings)—illustrate that all engaged providers have their internal processes sufficiently optimized to meet the overall targets in terms of time and quality.

Metric 11 (percentage variance in the imaging budget) illustrates a strict and successful financial governance of providers on a study level. Metric 12 (turnover of key project staff) illustrates that the selection process of providers with qualified available staff is successfully supporting a stable and reliable business partnership. Stability of staff positively influences most study management metrics, including immediate reaction on any issue during study conduct.

Metric 14 (percentage difference in expected versus actual reads) shows that the review process in CNS and inflammation studies is organized effectively, whereas the reading content and adjudication process needs adjustment and falls therefore clearly into group 2. Oncology studies need improved read planning.

Metric 16 (independent review charter lifecycle) is important, since it reflects the growing number of complex imaging studies that are engaging more than 1 imaging technique and

Table 20. Roche metric 16: imaging charter development lifecycle.

Metric #	Metric Description	Target	Clarification	ALL					CNS			ONCOLOGY			INFLAMMATION		
				Mean	SD	min	max	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
16	Independent review charter	42 d	1) Number of calendar days to produce first draft independent review charter - from "Date Contract Finalized and Signed" to "Date first draft Charter"	131.1	225.9	−102	942	42	547.0	324.1	6	68.8	105.6	30	26.8	68.3	6
		60 d	2) Number of calendar days to finalize independent review charter	113.7	140.5	1	624	44	186.8	249.5	6	85.2	91.8	33	214.4	198.8	5
		4 drafts	3) Number of versions pre finalization	2.1	2.6	1	17	43	3.7	6.5	6	1.8	1.1	32	2.1	0.7	5
		0 versions	4) Number of versions post finalization	0.6	0.9	0	2	13	—	—	—	0.5	0.8	12	2.0	—	1
		60 d	5) Number of calendar days to finalize reading database - from "Date Charter final" to "Reading Database Final"	133.5	131.8	5	585	22	—	—	—	135.8	141.1	19	119.3	53.3	3

Table 21. Roche metric 17: changes to imaging site manual.

Metric #	Metric Description	Target	ALL					CNS			ONCOLOGY			INFLAMMATION		
			Mean	SD	min	max	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
17	Number of image acquisition technique-related amendments per modality per protocol	≤1/modality/protocol	0.3	1.0	0	4	30	0.1	0.4	7	0.0	0.0	19	2.0	2.3	4

Table 22. Roche metric 18: financial health of preferred providers.

Metric #	Metric Description	Target	Clarification	ALL					CNS			ONCOLOGY			INFLAMMATION		
				Mean	SD	min	max	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
18	Financial Health	<0.5%	Bancrupcy prediction score % 2X per year - information from 2.1 of the RFI to be provided twice a year	0.0%	0.0%	0.0%	0.0%	12	—	—	—	0.0%	0.0%	11	0.0%	—	1

Table 23. Roche metric 19: preferred provider budget forecasting for studies.

Metric #	Metric Description	Target	Clarification	ALL					CNS			ONCOLOGY			INFLAMMATION		
				Mean	SD	min	max	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
19	Accurate quarterly forecasting	<3%	% difference	1.4%	3.7%	0.0%	16.0%	19	—	—	—	1.2%	4.4%	13	1.7%	1.5%	6

are asking for non-standard-of-care and quantitative imaging procedures. Imaging charters are a critical document, and it takes too long in our opinion to draft the charter and get final revisions done. Roche is working on standardizing imaging procedures to make the overall process easier for clinical sites as well as providers. However, this is difficult to achieve for complex imaging techniques in early clinical trials. At the same time, Roche teams need to provide input faster.

Group 2: Metrics That Need Refinement

The results obtained for metric 2 (days from image acquisition at site to image receipt at provider) have prompted Roche to encourage electronic image data transfers. However, this metric is misleading with regard to, for example, some oncology studies where all reading is done at the end of the study and there is no incentive for both the site and the provider to organize faster transfer of data. Therefore, it might be helpful to adjust the metric. Selected studies show that time-critical data transfers are achievable within 1 to 2 days. To achieve the overall target, good communication with clinical sites is essential. Metric 4 (days from image receipt to being ready for independent review) is influenced by batch readings as well. For oncology studies, where it takes a mean 33.7 days, this metric hints to internal provider process-improvement needs. This metric might, however, need an adjustment of target as well.

Metric 5 (days from when the image is designated for review to completion of the review) shows a well-organized process at providers for standard reading (RECIST, Cheson, etc). There is an adjustment needed for the more scientific input required for interpretation of demanding quantitative analysis processes, as used now in many early-phase trials, and for which the metric currently does not allow for.

Metric 10 (days from imaging study award to contract signature) shows the result of the interaction between Roche and providers and not so much discrete provider performance. Since many delays are caused by Roche's internal procedures, either the metric target needs to be adjusted, or the metric needs to be refined in addition to internal Roche changes.

Metric 13 (variance adjudication expected vs actual) shows that there is more work to be done on the forecasted versus actual adjudication rate. Although a number of publications[6,7] on adjudication percentage can be expected per response criteria used, it is still very variable per study. Despite all discussion currently ongoing, there is no other quality-of-reporting mechanism accepted by regulatory agencies that could therefore be used by the pharmaceutical industry. In comparing different studies with the same decision criteria and comparable settings, it is still not clear why there is such a high variability in reader assessment. This variability is even more pronounced for new decision criteria, such as RANO. Quality of reader selection and training content is therefore critical for a successful reading process. Roche and its providers have to work more closely and more stringently on this topic.

As a first step, all imaging charters have to include a clear definition of adjudication decision point (eg, time to progression) and target adjudication rate that is agreed at the beginning of the study. Therefore, as soon as adjudication rate comes close to the agreed target, study teams can react immediately. Metric 14 measures the handling of the reading process by comparing the expected amount of images needing to be read versus actual reads, and results show that most providers have this process set up well.

Metric 15 (inter- and intrareader variability) is important, as it provides data on the quality of the reading process. This has not been well reported by providers to date, and in future this will be integrated with metric 13 and adapted to ensure that it is reported accurately per study and overall by adding several submetrics: the expected adjudication rate, the actual adjudication rate in each quarter, the cumulative adjudication rate, and the actual intrareader variability for each quarter.

Group 3: Metrics That Are Discarded

Metric 18 (provider financial health) and metric 19 (quarterly forecasting) were introduced to see how good a provider was at forecasting the revenue flow for a specific study and the financial health of a given provider company. The forecasting was not widely used by study teams within Roche for imaging, so there is no need to report this. Financial health is also addressed separately by semiannual questionnaires and financial reports from the companies, and so there is no need for Roche to cover it as a metric.

The utility of metric 17 (amendments to site imaging protocol) is questionable, as most studies currently have no changes to the acquisition protocol once final.

Conclusions

The implementation of metrics has been useful in establishing good governance of providers to Roche. On one hand, it allows performance measurement for a specific provider for all studies as a mean and in individual studies; on the other, cross-provider performance as a whole and disease-area-specific results can be evaluated. It has also been of use in identifying areas of concern and in responding either to put processes in place to improve the specific metric or to work with individual providers to identify the cause for a metric being out of target and to put steps in place to improve.

As a consequence of these metrics, both Roche and its providers changed their specific processes to ensure better and more efficient ways of working together—specifically,

- implementing a new contractual process, reducing the time required to a mean of 24 days (from a mean of 74 days before implementation),
- currently working to implement electronic media capture for all studies to reduce the time from acquisition to receipt at site, and
- currently working to reduce charter development times and increase charter quality.

This has proven the effectiveness of most metrics in improving the relationship with providers.

However, the implementation of the Roche metrics has shown that each provider interprets the MCC metrics in its own way, and the findings from the first 4 quarters have demonstrated the need for clarification to the providers in terms of understanding and reporting each metric to Roche in a consistent and comparable way. This matches our understanding of the use of such metrics constituting a "living process." This also applies to Roche adjustments, such as metrics 18 and 19, which do not show value and were therefore deleted from the reporting list. Other metrics have had their specific targets adjusted (eg, time for contract signature has been increased from 14 days to 30 days), and this shows how experience is directly incorporated into the metrics package. The metrics again illustrate the importance of adjudication rate and how difficult it is to accurately forecast and implement on studies (metrics 13 and 15). This requires further investigation. With these exclusions, clarifications, and adjustments, Roche used this metrics scheme to the end of 2013, with rereview in the near future if necessary.

These metrics provide a snapshot of provider performance and identify areas for discussion as well as areas potentially requiring an audit to investigate further. However, metrics should not be used as the sole indication of performance, as there are also less tangible aspects (eg, team chemistry, how effective the project manager is, performance on individual studies) that need to be considered in an overall performance review.

The metric structure described herein has provided a basis for both Roche and the providers to work collaboratively and achieve benefits as a result of highlighting areas to become more efficient and delivery to expectations. The MCC metrics constituted the useful backbone of the Roche metric system. The majority of the metrics will continue to be used, and some will be refined to get a more granular understanding of the underlying process steps. Disease-specific target definitions and trial-phase-specific adjustments are also under discussion.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

1. Booz & Co Inc. 2012 healthcare pharmaceutical industry perspective. http://www.booz.com/global/home/what-we-think/reports-white-papers/article-display/2012-healthcare-pharmaceuticals-industry-perspective. Published December 7, 2011. Accessed January 24, 2014.
2. About.com. Key performance indicators: how an organization defines and measures progress toward its goals. http://management.about.com/cs/generalmanagement/a/keyperfindic.htm. Accessed January 24, 2014.
3. Metrics Champion Consortium. Building partnerships around standardized performance metrics. http://www.metricschampion.org/wp-content/uploads/2013/09/MCC_Imaging_PM_Handout_JUNE2013.pdf. Published June 2013. Accessed March 4, 2014.
4. Sullivan LB. Peak performance. http://www.samedanltd.com/magazine/13/issue/153/article/2943. Accessed January 24, 2014.
5. Quantitative Imaging Biomarkers Alliance. http://qibawiki.rsna.org/index.php?title=Main_Page. Accessed January 24, 2014.
6. Granger CB, Vogel V, Cummings SR, et al. Do we need to adjudicate major clinical events? *Clin Trials*. 2008;5:56-60.
7. Walovitch R, Memisoglu A, Walsh D. Adjudication in oncology trials. http://www.wcclinical.com/docs/whitepapers/200912-adjudication-in-oncology-trials.pdf. Accessed January 24, 2014.